

Statistical Research Article Critique

| | |
|-----|---|
| 著者 | Phillip Rowles |
| 雑誌名 | dialogos |
| 号 | 7 |
| ページ | 101-111 |
| 発行年 | 2007-03 |
| URL | http://id.nii.ac.jp/1060/00004990/ |



Statistical Research Article Critique

Phillip Rowles

There has been scarce research into learner autonomous recording of vocabulary and the subsequent learning strategies used by independent learners. As this topic has been neglected and is of interest to me, I perceive it as a potential future research area.

The purpose of this paper is to critique a survey research article written by Leeke and Shaw (2000) entitled, "Learners' independent records of vocabulary," published in *System* journal. This study was chosen because it was one of the very few survey research articles that have been written on this often disregarded research area.

The structural framework adopted in this paper consists of two main parts: first, a summary, and second, a critique, which is brought together in a conclusion. The critique, which is the main focus of this paper, examines the strengths and weaknesses of the study. The specific focus of the critique is on the appropriateness, adequacy, and interpretations of the statistics, also including the organization and quality of the study.

Summary

Leeke and Shaw's (2000) study entitled, "Learners' independent records of vocabulary," is a survey research article that includes both questionnaires and interviews. The study was divided into seven main parts: Background, Wordlists and the Vocabulary Acquisition/Retention Literature, Aim, Method, Results, Discussions, and Conclusions. In addition, there were two Appendices (A and B) and References

(from 23 sources) included. This summary will condense the major points of the seven main parts of the study in sequential order.

Background

Leeke and Shaw (2000) note that foreign language (FL) learners autonomous vocabulary needs depend on the subject they are studying, and individual learning styles. Lawson and Hogben (1996, as cited in Leeke & Shaw, 2000) concluded that FL learners when encountering unknown vocabulary, often used procedures such as writing the word and meanings into: 1) the margins of text pages, 2) word lists, or 3) word cards. Leeke and Shaw's (2000) study aimed at determining whether adult English as a Foreign Language (EFL) postgraduate learners at British universities followed the three procedures outlined above by Lawson and Hogben, (1996, as cited in Leeke & Shaw, 2000) or if they followed other procedures.

Wordlists and the Vocabulary Acquisition/Retention Literature

Leeke and Shaw (2000) state that several issues of second language (L2) vocabulary acquisition are pertinent. These are consciousness-raising, processing, lexis, psychological factors, multi-item word units, needs and self-awareness of adult EFL autonomous learners. There had been research into these related areas of L2 vocabulary acquisition, however, Leeke and Shaw (2000) noted there had been a stated lack of research into adult EFL learner word listing routines.

Aim

Leeke and Shaw's (2000) study aims were exploratory because of the stated scarcity of related published research. Their study covered only one sample, in one situation, therefore, the results could only show expected varieties and factors that might have been functioning. A literature review inspired them to investigate five

aspects of autonomous word list use: 1) How long did participants use word lists? 2) Why did they continue, or give up? 3) Were mnemonic devices used, or adopted? 4) What register (that is, technical, semi-technical, or colloquial) and medium (that is, television, classes, novels, etc.) did the words come from, and how did these relate to intended use? 5) How many multiword units were recorded, and how did this relate to the list purpose?

Method

Leeke and Shaw's (2000) study was carried out at Newcastle University, UK, over two years with different participants. Both groups of participants were adult EFL university "overseas postgraduate" learners. In Year One, a multiple-choice questionnaire was administered to 121 participants. In Year Two, semi-structured interviews were conducted with 54 participants. The questionnaire and interview questions were displayed in Appendices A and B respectively, of Leeke and Shaw's (2000) study.

Results

Leeke and Shaw's (2000) study results answered six questions. First, the proportion of learners who used word lists, indicated that in the questionnaire, 81 of the 121 participants recorded them, and in the interviews, 19 of the 54 recorded them. The interviewees were subjected to a visual inspection of their word lists in order to reconfirm their usage. Second, why they gave up, or continued, was affected by three factors: 1) individual learning style, 2) level of learning, and 3) extrinsic motivation. No clear separation between questionnaire and interview results was given. Third, mnemonic devices utilized in questionnaires indicated that the majority, 28 of the 81 list makers recorded translations only, while the majority of interviewees, 16 of the 19 list makers recorded translations. An interesting note was made of the three

Japanese interviewees who kept Japanese-manufactured pre-printed word notebooks with headings labeled on top of blank columns. These commonplace notebooks in Japan are apparently a rarity in the United Kingdom. Fourth, the ordering of lists from questionnaire results indicated that the majority, 27 of the 81 list makers, used random order, and the interviewees were stated as conducting random ordering at an even higher proportion. However, this interviewee proportion was not given. Fifth, the register, medium, and relation to intended use from questionnaire results indicate that the majority, 30 of the 79 list makers, recorded technical vocabulary, while interviewees results were not given. As for medium, or source, interviewee results indicated that the majority, 11 of the 19 list makers sourced from specialist reading. The questionnaire did not include an item, or question, related to this topic. Intended use interviewee results indicated two purposes: one was direct for production, and included a many multi-item words and infrequent translations, while lists for general language improvement were prone to include single words and translations. The questionnaire did not include an item, or question, related to this topic. Sixth, the extent of multi-item word use was indicated by interview results that the majority, 11 of the 19 list makers, recorded individual words. The questionnaire did not include an item, or question, related to this topic.

Discussion

The previously published literature on word listing routines indicated two contributing factors: 1) psychological, and 2) linguistic. Leeke and Shaw's (2000) study indicated an additional two factors: 1) personal and motivational, and 2) learning level and context issues. All of these four factors, along with other factors, were presented visually in Figure 2 on page 286 of Leeke and Shaw's (2000) study.

Conclusions

Leeke and Shaw (2000) concluded there were differences between theoretical and actual vocabulary listing strategies, and between teacher-led and autonomously crafted word lists. Leeke and Shaw (2000) reiterated that this study focused on actual examples drawn from autonomous participants, separating this study from other studies focusing on theoretical, teacher-led word list research. This more practical approach was deemed to be more useful for giving realistic advice to other adult EFL learners. Stated effective advice suggested that guidance should also take into account individual learner differences and needs, to boost motivation and allow all the factors in Figure 2 (p. 286): that is, individual factors, purpose, colleague examples, language beliefs, previous training, learning beliefs, learning stage/context, and time constraints, to operate more productively.

Critique

The following critique examines the strengths and weaknesses of the study and focuses on appropriateness, adequacy, and interpretations of the statistics, as well as the organization and quality of the report.

Appropriateness

A primary examination of the genre or place of publication is revealing. *System* journal's description from their website states that it is an international journal covering aspects of educational technology and applied linguistics. Of particular note was the Statistical Research 6 description, "The journal serves as a vehicle of expression for colleagues in developing countries," from *System's* website. We might assume that *System* journal is not in the same league as other higher-caliber journals such as *Language Learning* or *Language Testing*. This may become evident as this critique unfolds. This is not to say that studies published in *System* journal

are irrelevant, it is merely pointing out that the studies have been published with a specific audience in mind. The focus might be categorized as less theoretical, or technically-based, with more emphasis on practical application. Therefore, the critique has been written taking this audience register taken into account. Overall, as it was indeed published by *System*, whose editorial board includes such L2 research luminaries as Dick Allwright (Lancaster University, UK) and Rod Ellis (University of Auckland, New Zealand), one would have to say this study was appropriate.

Unfortunately, *System's* website does not detail the process of blind-reviewing, or the reviewing process at all, so this is an unknown quantity.

The appropriateness of the statistics in the study is suitable, in that descriptive statistics are employed, and the survey research methodology is exploratory in nature.

This exploratory style matches the stated lack of related previously published research in this field of adult learner autonomous vocabulary list use, and strategies.

Adequacy

The lack of previously published research prior to 2000, when this study was published may be true. However, while there may be a lack of literature in these fields, other published research (pre-2000) certainly exists and was not cited by Leeke and Shaw (2000), for example, Pino-Silva (1993), Griffin and Harley (1996), Schmitt (1997), and Statistical Research 7 Sokmen (1997). Expanding the review of previously published literature may have added greater objectivity to Leeke and Shaw's (2000) study.

The adequacy of the collection of statistics in Leeke and Shaw's (2000) study was good as the descriptive statistics were mostly presented as frequencies. However, taking the next step, that is, analyzing the survey data statistically, by for example, using the simplest form of comparing frequencies, that is, the Chi-square (χ^2)

statistic (Brown, 2001), was not attempted, and therefore not adequate.

Interpretations

In regards to Chi-square, even though it is the simplest form of comparing frequencies, it is plagued by many problems and pitfalls (Woods, Fletcher & Hughes, 1986). These problems and pitfalls include small expected frequencies, the two-by-two contingency table, independence of observations, testing several tables from the same study, and the use of percentages. Therefore, it is recommended that Leeke and Shaw (2000) should have designed a questionnaire that consisted of selected-response items using a Likert scale (Brown & Rodgers, 2002). From here there are many more options.

As Leeke and Shaw alluded to factors visually in Figure 2 on page 286, they could have analyzed this statistically instead, by performing a factor analysis of the recommended selected-item responses and Likert scale-based questionnaire. In the end, no statistical analysis of the closed-ended items in the questionnaire was conducted. This was a basic research design flaw. Another major interpretation problem was the reference to inferential statistics in some of the questions, even though only descriptive statistics were collected. This is illustrated by the questionnaires examining 121 participants and the Statistical Research 8 interviews examining 54 participants at one British university (Newcastle University).

However, the first question in the results, 5.1, states, "What proportion of overseas postgraduates at British universities make word lists" (Leeke and Shaw, 2000, p. 275).

The qualitative analysis of the interview results was sometimes reported as frequencies, and sometimes individual participants responses were reported. This seemed to narrow the study down to a micro level, or disjointed case studies, where different participants were targeted according to the question. Leeke and Shaw

(2000) hint at this themselves in the Method section, "Our selection of interviewees was not systematically random, so what we have is a series of case studies rather than a valid sample" (p. 275). As a result there seemed to be little consistency or coherence in the qualitative analysis of the interviews.

Organization

Leeke and Shaw's (2000) study had an overall strength in that it included both questionnaires and interviews as examples of survey research. The two major problems associated with this selection were that first, the questionnaires were conducted in the first year, and the interviews were conducted in the second year, using different samples. While there is nothing inherently wrong with using two samples, the use of one sample in a repeated-measures type design may have been more illuminating. Second, the questionnaires were administered and a year later, the interviews were conducted. This was the reverse of the sequence recommended by Brown (2001) who saw these two ways as complementary but "interviews are more suitable for exploring what the questions are and questionnaires are more suitable for answering those questions" (p. 79). As Leeke and Shaw (2000) stated that this study was exploratory in nature, using interviews first before formulating concrete items for the questionnaire would have been better.

Quality

Leeke and Shaw's (2000) study quality is questionable as often there were unmentioned critical details. Two missing details that come to mind were first, there was no reporting of piloting taking place. This essential first step was missing in the study. Brown (2001) stated that the purpose of piloting a questionnaire was to see the type of answers that would be obtained and make appropriate changes to avoid any problems. At the very least, Leeke and Shaw (2000) could have tried one of two

procedures. Piloted before the first questionnaire administration, or treating the Year One questionnaire as a pilot and then giving it to the interviewees in Year Two. The second missing detail was who conducted the interviews. Was it the two writers of this study, Leeke and Shaw? If it was Leeke and Shaw, did they pilot the interviews to obtain a measure of inter rater reliability? Unfortunately, this was not reported in the study.

Another area of concern was the level of respect, trust and rapport developed with the participants. The reported data of the interviews only acknowledged as vocabulary list makers those participants who brought physical evidence of their word lists. If an interviewee said they made lists but did not bring any proof, they were not rated as a vocabulary list maker. The Discussion stated that a "little more than a third of our subjects continued to make vocabulary lists when they were learning independently in an L2 environment" (Leeke and Shaw, 2000, p. 283). While this was certainly true for the interviewees, it is misleading, as the Discussion does not report that two-thirds (or 81 of 121) of the questionnaire participants indicated the use of word lists. Belief without proof seems to be an issue here, despite the fact that the first item in the questionnaire asked if vocabulary records were made. The unanswered question is, why would researchers include such items if they were not going to believe the answers?

In both the questionnaire and the interview, the structure was inappropriate. In the case of the questionnaire, it was structured in such a way that word list makers were asked eight items, while non-word list makers were asked only two items. As the non-word list makers (40 of 121 participants) would have finished much quicker than the other two-thirds of questionnaire respondents, this represents another basic design flaw. Rewriting the questionnaire, as was mentioned previously, to consist of selected-response items and Likert scales (Brown and Rodgers, 2002) would eliminate this design flaw and have the ideal situation of every participant responding

to all items. Similarly, in the interviews, word list makers were asked a total of 13 items, while non-list makers (including those who said they were word list makers but did not bring proof) were only asked 6 items. Possibly in the interviews case it is not such a major problem as non-word list makers interviews would probably only be finished much quicker.

Conclusion

Leeke and Shaw's (2000) study examined two elite groups, or two samples, of participants: adult postgraduate EFL learners at Newcastle university a year apart using questionnaires and interviews. As these two samples were represented by 121 and 54 participants in the questionnaires and interviews respectively, the results are not generalizable to a larger population by the use of inferential statistics as alluded to by Leeke and Shaw (2000).

In conclusion, despite many weaknesses reported in this critique, Leeke and Shaw's (2000) study seemed well-suited to *System* journal, as it was not heavily theoretically-based, but rather more practical in nature. This learner-centered focus was refreshing and realistic for the intended audience of instructors. Analyzing the data statistically would have added more weight to their argument, and would probably have been perceived as more professional and rigorous, possibly enabling Leeke and Shaw to publish in another journal.

Despite these concerns, the study was published in *System*, which has a related website that attests it is an international journal of educational technology and applied linguistics. Overall, Leeke and Shaw (2000) must be congratulated on achieving the admirable goal of publishing a study in an international journal.

References

- Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.
- Brown, J.D. and Rodgers, T.S. (2002). *Doing second language research*. Oxford: Oxford University Press.
- Griffin, G. & Harley, T. (1996). List Learning of Second Language Vocabulary, *Applied Psycholinguistics*, 17, 443-460.
- Pino-Silva, J. (1993). Untutored Vocabulary Acquisition and L2 Reading Ability, *Reading in a Foreign Language*, 9, 845-857.
- Schmitt, N. (1997). Vocabulary Learning Strategies, in Schmitt, N. & McCarthy, M., editors (1997) *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press.
- Sokmen (1997). Current Trends in Teaching Second Language Vocabulary, in Schmitt, N. and McCarthy, M., editors (1997) *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press.
- System journal website (2006). Retrieved June 10, 2006, from:
http://www.elsevier.com/wps/find/journaldescription.cws_home/335/description#description
- Woods, A., Fletcher, P., and Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.